

**MEDIN Standards Meeting
London 1st December 2008
10.30 – 16.00**

Minutes

Mark Charlesworth (MEDIN) – WS Lead
Steve Gontarek (SAMS)
Colin Graham (BGS)
Dave Cotton (MEDIN)
Roy Lowry (BODC)
James Dargie (CCW)
Chris Hill (GeoData)
Jason Sadler (GeoData)
Steve Wilkinson (JNCC) - Chair
Daniel Lear (MBA)
James Rapaport (SeaZone)
Becky Seely (MBA)
Roger Coggan (CEFAS)

1) Introduction and Welcome

All were welcomed to the first meeting of the newly formed MEDIN standards group. Apologies were received from Kieran Millard.

2) Outline of objectives for the day (MEDIN)

MC gave an outline of the objectives for the day, stating that any technical terms should be explained clearly for the non experts.

3) Terms of Reference for Standards WG

The terms of reference for the group were taken from the objectives and rational statements in the MEDIN Work Programme agreed at the sponsors board, and the previous terms of reference from the MDIP Standards WG. These were then updated to reflect new working structures and circulated to the group. The new terms of reference were generally agreed however further explicit statements on the need to consider metadata and data may be required.

AP. MC to discuss with SW and edit as necessary then circulate to group for final sign off.

4) Discovery Metadata Standard and Tool

a) Metadata requirements

JNCC and GeoData gave updates of relevant initial finding from the portal review and archiving data across DACs projects respectively. The consistent level of metadata (granularity) was considered key to providing a useful resource and also agreement on hierarchy terms (programme, project, series, survey). The choice of keyword list was also deemed critical. For mandatory fields a balance must be struck between what can be expected submitters to fill in and what is needed to discover the data. The NERC data grid and OGC WCS (Open Geographic Consortium Web Coverage Services) structures have been considered for archiving data across DACs. A hybrid approach was proposed which would allow flexibility on where the streams of data originate for presentation to the portal.

b) Metadata standard

The proposed new metadata standard was presented. Specific points were as follows:

There was some replication in elements which needed addressing or further explicit guidance. These included Originating Controlled Vocabulary, Conformity, and Unique Resource Identifier.

Element 'Title'. A common format was suggested akin to that already in use by the NBN. Where possible this could be built from existing metadata elements.

Element 'Abstract'. Generic guidance should be used to ensure consistency.

AP. CG, RL to provide examples to DL/BS.

Element 'Topic Category'. This should follow ISO19115. Further definition could be provided using the Keyword.

Element 'Keyword'. A number of lists exist including SeaDatNet, MESH and GCMD Science. It was agreed that a comparison of the SeaDataNet keyword lists to other should be made and preferably changes made to the SeDataNet list. The SeaDataNet lists are hierarchical at 4 levels – the levels most appropriate were parameter disciplines (P081) or parameter groups (P031). Some consideration of the INSPRE themes should also be made.

AP. MC to circulate link to SeaDataNet parameter thesaurus and BODC vocab server.

Element 'Originator'. Should be made mandatory due to the role of DACs in managing data on behalf of others. Further consideration of how to deal with organizational name changes is required.

AP. DL to consider how changes in organizational names would be handled.

Element 'Lineage'. Unsure if this element referred to the methods used to capture the data or the process history or both. Guidance based on INSPIRE definition is required. Should be made mandatory.

Element 'Extent' It was agreed that if possible this should be generated from the bounding coordinates and tools to develop this were being build under a project 'SeaVox' which BODC are involved in. It was recognized that number of sea regions exist and it was not agreed which to use.

AP. All to send any definitions of sea-regions to DL for consideration.

Elements 'Vertical Extent' and 'Spatial Reference System'. Explicit guidance on what these elements were and their relationship to other elements is required.

Element 'Unique Resource Identifier'. Agreed that MEDIN should not attempt to issue DOIs and the field should not be mandatory. However, the DACs should consider implementing unique keys to the metadata records that they submit to the portal.

AP. MC to raise with DACS WG.

Elements Metadata Standard Name and Version. Should be included in the schema.

b) Metadata creation tool

DASSH gave a presentation of the current Fish DAC metadata tool which is able to pre-fill fields on the basis of the registered individual who is entering the information. All agreed it was encouraging, although it would also be worth considering a SeaDatNet tool Mercardo which can handle both form input and also SQL tools for direct import which could potentially be very useful particularly for transferring data between organizations and initiatives.

Some guidance is required on what granularity was expected. Generally thoughts included:

- correct level is a dataset such as cruise, survey or a set of repeat observations with a common purpose,
- a specifically funded piece of work
- the dataset should be able to be easily extracted from a database for a 3rd party
- if you get the result every time you search by a combination of time, location and parameter then it is probably too coarse

AP. DL to work with MC and CG to draw up guidance using the above as a starting point.

The potential to employ a specific person to help the community to draw up metadata records was considered which MEDIN had already discussed. GeoData provide training via workshops which may be a model to use. Some concern if somebody could have the experience across all domains.

The other point of discussion was who writes the metadata record and the flow of data and metadata from contractor to organization to DAC to portal.

AP. DC to consider above cross cutting issue and propose strategy to address.

5) Data Product Specifications

MC gave a brief presentation to set the scene. It was felt that to provide data specifications akin to those being produced by INSPIRE within a large generalized model was not required and that the work at this level should concentrate on providing 'best practice documents'. These should specify the information that should be collected and held with a dataset and potentially suggest controlled vocabs to be used and would be domain specific. These could then be used by the DACs and organizations to append to contracts. In the first instance current best practice documents should be gathered in a library to identify gaps and then coordinated work with CCW and DASSH would progress building new data specifications. Controlled vocabularies would not be considered separately to Best Practice Documents.

AP. MC to collect library of 'best practice documents' by domain and put on secured area (i.e. not publicly available) of MEDIN website by 1st Feb.

AP. All to send current 'best practice documents' to MC for collation.

6) World Register of Marine Species (WoRMS)

WoRMS is a species register that MEDIN are helping partners become WoRMS compliant by coordinating UK organisations to register new species and do comparative checks.

7) Standards for exchanging GIS, charts etc.

It was highlighted that the group would have to address standards for GIS at some point and if there was any timescales. It was recognized that this is something that is required by next summer for the MPA project and further guidance on the discovery metadata standard should be provided. Partners with experience of using GIS and generating GIS metadata were asked if they would be interested in providing some guidance notes.

AP. All consider if their organisations could provide guidance notes for generating metadata for GIS layers and contact MC if so.

8) Signing off and dissemination of standards

It was agreed that any new standard should be implemented in the following manner:

1. drafted by experts in the domain
2. key feedback sought from 2-3 designated people and distributed to MEDIN standards WG and any relevant DACs
3. revision following 2)
4. sent to all MEDIN partners in draft form (min 2 weeks for comments)
5. revision following 4)
6. ratification by MEDIN Executive Team
7. publication on MEDIN website

Emphasis would also be on partners to distribute and enforce any new standard in their own organisations. For the metadata discovery standard it was agreed that the schema, guidelines and tools should be sent to partners on 1st Feb with 2 weeks for comments and then following revisions will be ratified by the MEDIN Executive Team on the 27th February.

9) Future working arrangements for group

- Meeting frequency – it was agreed to meet every 6 months however the next meeting would be in April to agree the work programme for 2009-2010
- Chair – it was agreed that Steve Wilkinson should remain chair to ensure continuity

- Participants – it was felt that the current participants had the required range of expertise and the group was of a reasonable size. Further involvement from NI could be sought.

AP. MC to discuss with NI representatives in February.

- Meeting Location – It was agreed that the group should have a range of meeting locations to ensure full participation. The next meeting would be in Edinburgh.

10) AOB